Modeling and Analytics (M&A)

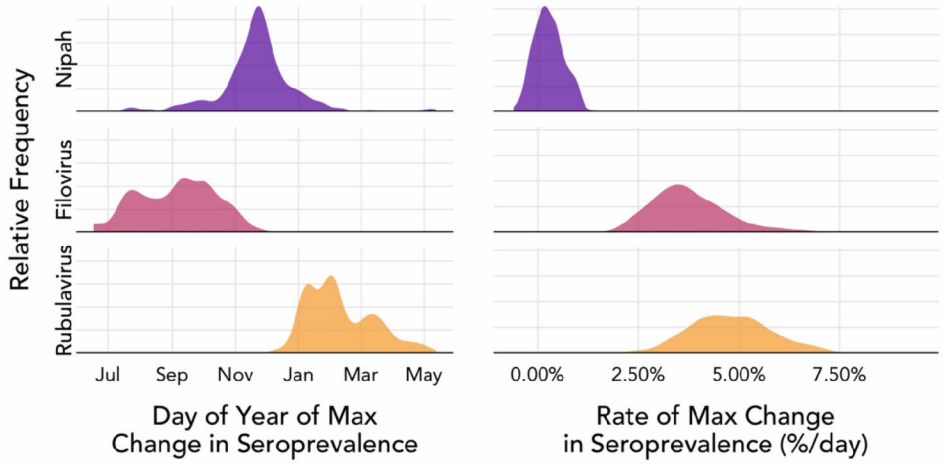
Highlights and successes

In the last six months, PREDICT M&A has focused on four major themes:

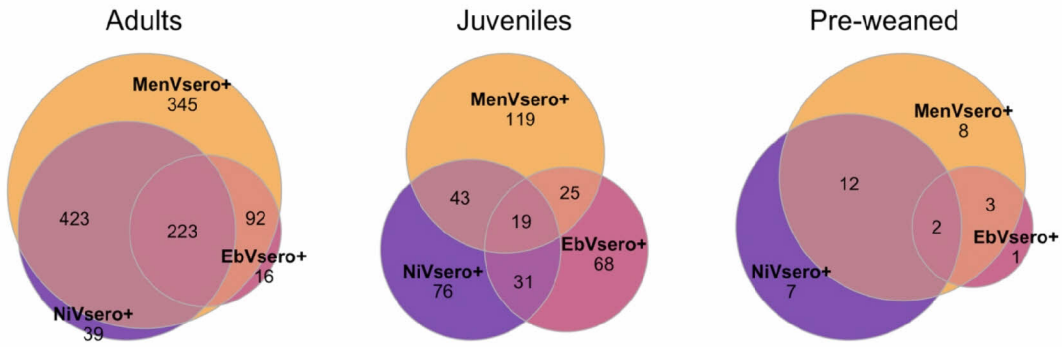
1) Demonstrating how we can **analyze PREDICT data** to increase the efficiency of surveillance programs and predictive power of our models
2) Intervention Modeling Projects ACross Teams (IMPACT) that help **test proposed broad-based intervention strategies that derive from PREDICT results**;
3) **Working with EPT partners to produce maps, models, and other products** that help them design their programs, test hypotheses, and hone their One Health strategic plans; and
4) Designing **new analytical strategies and open-source tools to analyze risk of disease emergence**.

1) Analyzing PREDICT data to increase efficiency of surveillance programs

The PREDICT M&A team conducted analyses to examine whether surveillance can be targeted seasonally to a period when the risk of disease emergence is greatest. We used longitudinal serological data from Bangladesh to identify co-circulation dynamics of Nipah, filovirus, and Rubulavirus in a bat population of the species *Pteropus medius*. We show that each virus has different periods/months when seroprevelance is significantly increasing or decreasing, and the calculated risk of viral shedding is greatest. We also analyzed these data to show that individual bats can be co-infected with multiple viruses. **Therefore, interventions to mitigate the spillover of viruses from this one fruit bat species will need to consider different periods of viral shedding, and a single intervention to mitigate human exposure may have the benefit of reducing spillover risk for multiple viruses.**

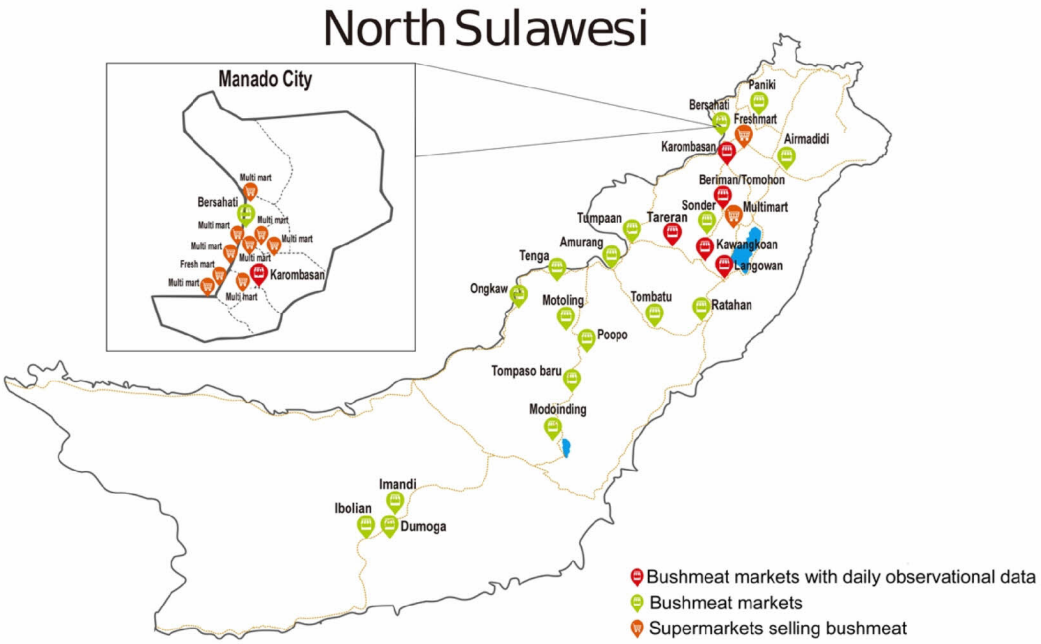


Analysis of serological data for three viruses to determine the timing and strength of viral circulation in juvenile bats from a longitudinal dataset from Bangladesh. Plots show periods of the year when the increase in seroprevalence is greatest (left), and the rate of change for each virus (right). Plots display the relative frequency from 1,000 generalized additive model runs.



Patterns of seropositivity to multiple viruses in fruit bats from Bangladesh. Venn diagrams for adult, juvenile, and pre-weaned (dependent) Pteropus medius *bats testing positive for antibodies against of Nipah virus, Filovirus, and Rubulavirus. Numbers under labels are counts of bats with only seroreactivity to those viruses, numbers in overlapping areas represent number of bats detected with multiple viruses or evidence of co-immunity.*

PREDICT is analyzing the specific risk of viral emergence across the wildlife market value chain. <u>Using data from bushmeat market surveys in North Sulawesi, Indonesia</u>, models are being developed to quantify the areas of greatest risk for bat- and rodent-borne zoonotic virus emergence. We identified markets that sell fresh and frozen bats, rats, wild pigs, and snakes. Daily observational data are being used to calculate the volume of wildlife sold monthly/annually and will feed in to IMPACT projects to assess the specific risk that these markets play in disease spillover and spread.

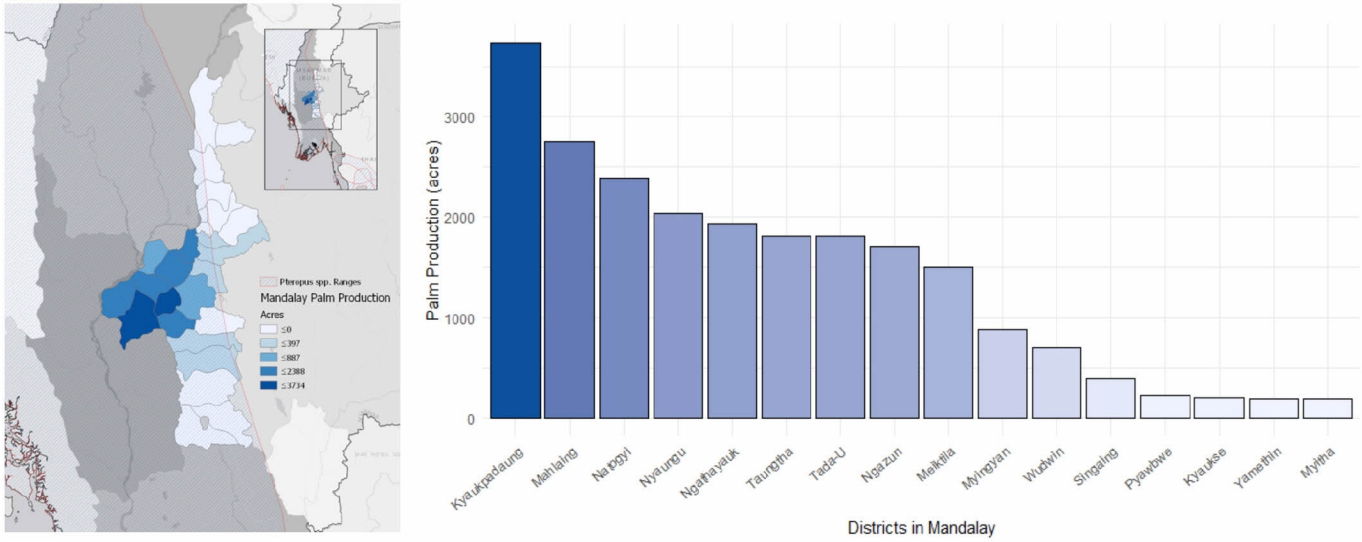


Map of the wildlife trade/value chain in North Sulawesi, Indonesia. Green markers indicate bushmeat markets surveyed by PREDICT, and red denotes markets where daily observational data were recently collected. Supermarkets where bushmeat is also available are marked in orange.

2) Analyses to test proposed broad-based intervention strategies

The PREDICT M&A, Surveillance, and Behavioral Risk teams continue to develop analyses to assess potential evidence-based risk interventions and define the boundaries within which interventions might prove successful. Over the last six months, significant progress has been made on 15 **IMPACT** projects (**I**ntervention **M**odeling **P**rojects **AC**ross **T**eams) focused on different risk factors or specific interfaces of disease emergence.

For example, one IMPACT project is examining the risk of bat-borne disease emergence at the **Bat-Date Palm Sap production interface.** While the risk of Nipah virus emergence from date palm sap consumption in Bangladesh is known, the risk of disease spillover via this interface in neighboring countries is not. PREDICT is collecting and analyzing date palm sap production data from Myanmar and incorporating it into models to map the ecological and geographic range of key bat species. Together, these combined agricultural and wildlife distribution models will identify potential sites of viral spillover where surveillance efforts in people and bats should be heightened.



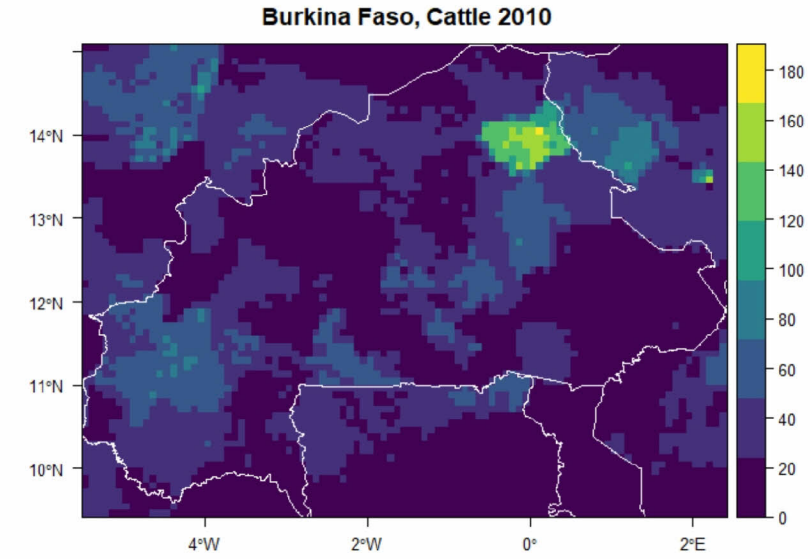
Date palm sap harvesting data from Myanmar by district in Mandalay being used to map the risk of bat-borne virus spillover risk.

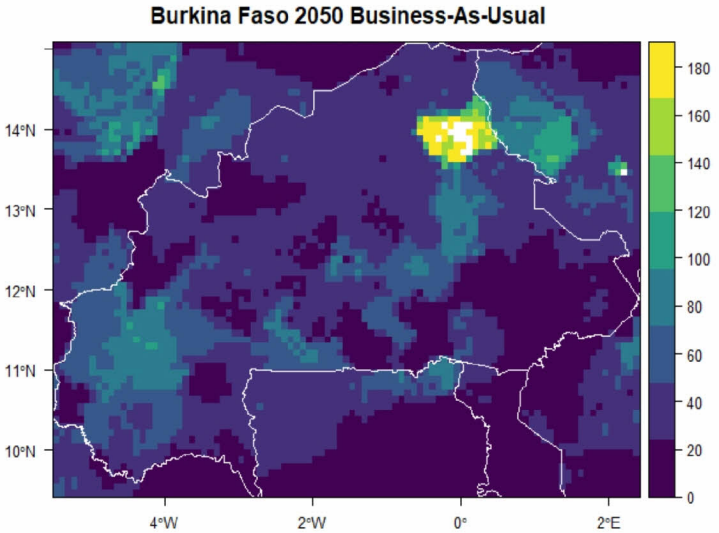
Over the remaining six months of the project, the M&A team will continue to coordinate across PREDICT, USAID EPT, and our in-country teams in modeling, capacity-building, and scientific communication efforts. Outputs from all analyses will be developed as

short, policy-relevant Emerging Disease Insight documents for distribution to key stakeholders and the public.

3) Working with EPT partners to produce maps, models and other products of value for their programs

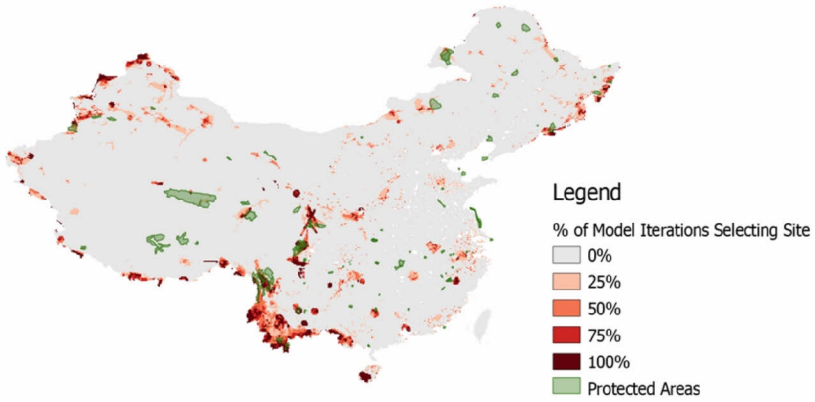
PREDICT continues to collaborate with FAO on the Africa Sustainable Livestock 2050 (ASL2050) project to model the risk of disease in the face of expanding livestock production in Africa. We produced geospatial maps of cattle density under three future scenarios of agricultural development: 1) Stratified societies, 2) Business as usual, and 3) Toward sustainability. Countries included in these forecasts: Burkina Faso, Uganda, Kenya, Nigeria, Egypt, and Ethiopia.





Forecasted changes in cattle density for Burkina Faso from 2010 to 2050 under a business-as-usual scenario. Maps produced by PREDICT as part of the Africa Sustainable Livestock 2050 (ASL2050) project in collaboration with FAO.

PREDICT's M&A team supported the development and strategic design of the Global Virome Project (GVP). The PREDICT team has designed a spatial modeling approach to identify priority sites for targeted wildlife sampling at a 10 x 10 km resolution in countries of relevance to the GVP. PREDICT developed specific maps for the Thai National Virome Project and the China Virome Project and presented these at the TNVP launch, as well as in high-level meetings with Chinese Government and US Embassy leaders in Beijing. These analyses will form the basis for the design of specific workplans in both countries during the rollout of their virome projects in 2019-20.

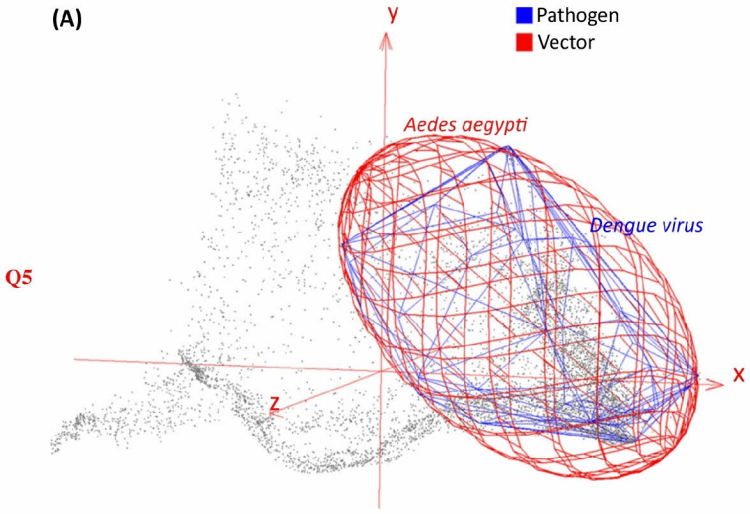


Optimizing sites for the China Virome Project, part of the GVP. Darker shaded areas on the map (red and dark red) are sites that were consistently identified by the model as the optimal areas to discover the greatest number of viruses using the fewest resources. Protected areas in China shown in green.

4) New analytical strategies and open-source tools to analyze risk of disease emergence

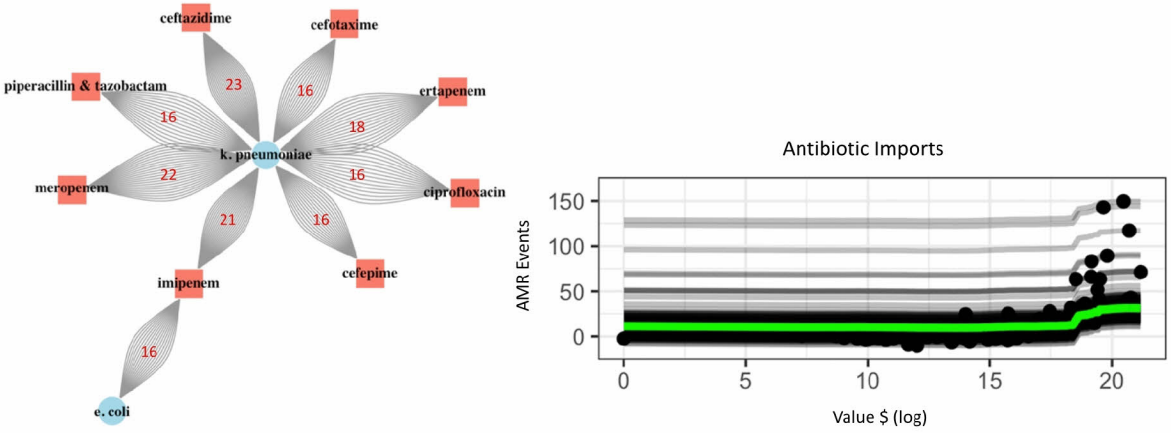
The PREDICT M&A team worked closely with the Information Management team to **continue developing new tools within the EIDITH R package to facilitate the cleaning and analysis of project data**. These open-source tools allow country and global teams to download their country-level PREDICT-2 data, clean datasets by automatically identifying outliers and fields with missing or incorrect data, and analyze patterns in the data using the statistical analysis program, R. Site characteristics, behavioral risk, animal, or testing data can then be manipulated in R to explore and visualize data from the project in near real-time, as it's entered into the database.

PREDICT developed a new theoretical framework for modeling the geography of disease transmission using biological properties of both hosts and parasites to produce reliable outputs of disease distributions. These analyses can be used to better map and help guide the surveillance and discovery of pathogens across the landscape.



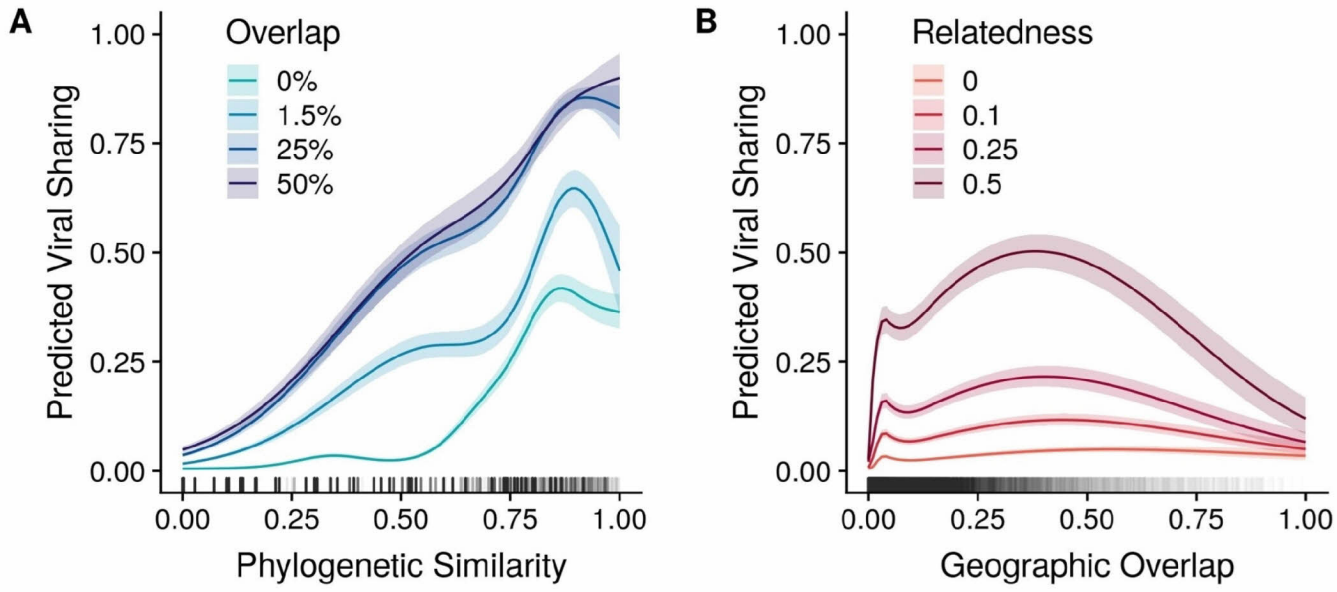
PREDICT developed a new theoretical model of disease distribution using ecological niche models. Examples above shows how dengue virus (blue polyhedron) distribution coincides with that of its vector, the mosquito Aedes aegypti (red ellipsoid). In Press, Trends in Ecology and Evolution.

PREDICT completed the most comprehensive database to date to analyze the drivers of antimicrobial resistance (AMR) disease emergence in people. This effort involved data extraction and data cleaning from over 24,000 scientific papers published from 2006-2017. PREDICT is currently using this database to develop an AMR "hotspots" map to show areas around the world where future AMR emergence is most likely. **Preliminary analysis** using Bayesian Additive Regression Tree (BART) models are underway to identify potential risk factors that explain the number of AMR events within a country. Variables include: research bias, percentage of agricultural land cover, livestock index, GDP, and % of GDP spent on healthcare, population, antibiotic imports, and antibiotic exports.



The most common drug-bacteria combinations in PREDICT's AMR database are displayed, the number of emergence events for each bacteria-drug combination are shown in red (left). Example of Bayesian Additive Regression Tree (BART) model output showing an increase in the per-country risk of a new AMR emergence event with increasing levels of antimicrobial drug imports for a given country (right).

Building off previously published PREDICT analyses to identify host and viral traits that predict zoonotic potential, we developed **new models to explain viral sharing between all mammal species**. Specifically, phylogenetic similarity and geographic overlap were used to predict a viral sharing network across all known mammals (~5000 species). These analyses found that more closely related species and those that overlap the most in space, are much more likely to share viruses. The relationship of these factors was non-linear, and the two variables interact so that phylogenetic similarity has a stronger effect on species that overlap more, and conversely overlap matters more when species are more closely related.



New PREDICT model to examine factors that determine viral sharing among all known mammal species. These plots show the interacting, non-linear effect of host species relatedness (left) and geographic overlap (right).

For more information

A full list of PREDICT M&A team products and outputs are included in the *Monitoring and Evaluation Appendix 1*. Emerging Disease Insights are available online: http://livescience.ecohealthalliance.org